

Education Policy Studies Division
Contact: Liam Goldrick, 202/624-5359
or lgoldrick@nga.org
December 9, 2002

Improving Teacher Evaluation to Improve Teaching Quality

Summary

Governors understand the importance of guaranteeing that every child has an effective teacher. Research shows that teacher quality affects student achievement more greatly than any other school-based variable.¹ The No Child Left Behind Act requires a “highly qualified” teacher in every classroom by the 2005–2006 school year and achievement gains by all students over time. These realities give policymakers a strong incentive to focus on preparing, recruiting, and retaining quality teachers as primary strategies to boost academic achievement.

As more states develop accountability and data systems capable of tracking learning gains by individual students, greater opportunities will arise to measure the value-added impact of teaching. Such an approach to assessment may help revolutionize teacher evaluation, which traditionally has been based on procedural reviews and infrequent classroom observation by often untrained and overtaxed school administrators. However, states can make important changes, even within the current accountability framework. By linking evaluation with academic standards for students and professional standards for educators, policymakers can transform teacher evaluation into a more effective tool for improving instructional practice and raising student achievement.

To improve teacher evaluation, Governors and other state policymakers should employ the following strategies:

- **Define teaching quality**—*States have defined academic standards for what every child needs to know. They also must clearly define what a highly qualified teacher needs to know and be able to do before they can purposefully construct a teacher evaluation policy.*
- **Focus evaluation policy on improving teaching practice**—*States should view evaluation as an informational tool to help administrators identify teachers who need additional or specialized assistance and to help individual teachers improve their instructional practices.*
- **Incorporate student learning into teacher evaluation**—*States should transform evaluation from a traditionally input-based process into an outcome-driven one. They should consider measurable student achievement as a principal outcome on which teachers are evaluated.*
- **Create professional accountability**—*Career ladders can provide states an opportunity to strengthen teacher evaluation policy and align it with performance-based teaching standards. Professional classifications (such as “beginning,” “mentor,” and “master” teacher) can also provide a framework through which to implement performance-based compensation.*
- **Train evaluators**—*Evaluators need preservice training opportunities to conduct more accurate and effective teacher assessments. Training might focus on skills such as analyzing effective teaching practice, determining a teacher’s impact on student learning, and providing leadership for professional development and remedial assistance.*
- **Broaden participation in evaluation design**—*Policymakers must reach out to all education stakeholders, including teachers and administrators, to design a teacher evaluation system. Educators and school officials must have confidence in and an understanding of evaluation—prior to and during implementation—to ensure its long-term sustainability.*

What Is Teacher Evaluation?

Evaluation is the process by which teachers are assessed professionally. Usually conducted by principals or school administrators, it may include classroom observation as well as verification of continuing education and professional-development activities. Evaluation is distinct from teacher licensure—a means of credentialing beginning teachers—which often requires completion of a basic-skills test and evidence of minimum competence. Evaluation seldom is integrated into or informs the re-licensure of veteran teachers.

Teacher evaluation typically has been designed as a personnel action, not as a tool for instructional improvement. Though evaluation serves as a mechanism for assessing job performance, in practice it is often cursory, subjective, and based upon insufficient observation. Moreover, it seldom results in the termination of truly poor educators. Charlotte Danielson, a development leader at Educational Testing Service, describes evaluation as “an activity that is done to teachers.”² Similarly, a Massachusetts education consortium calls it “a task that teachers endure, a task conducted by already overextended school administrators.”³

Current state and district policies give school leaders broad latitude in designing the evaluation process. Principals may evaluate teachers annually, but only observe them every two years to four years, awarding tenure based on evaluations that are largely disconnected from student learning and that do not help improve teaching practice.⁴ Increasingly, however, policymakers are scrapping this approach in favor of performance-based teacher evaluations.

A purposeful evaluation system measures teaching *outcomes*, not simply teaching *behavior*. Evaluations that are well-designed and integrated with curriculum and professional standards can accomplish more than assuring basic competence. They can help states and districts measure the effectiveness of teachers at various points in their careers, identify highly skilled teachers, offer specific recommendations to improve teaching, inform professional development, and demonstrate accountability for student achievement. State policymakers should treat teacher evaluation as an integrated component of a comprehensive strategy to improve overall teaching quality.

State Actions to Improve Teacher Evaluation

The No Child Left Behind Act provides Governors and other state policymakers an opportunity to enact or amend laws and regulations governing teacher evaluation, alongside other required reforms. The inclusion of evaluation into broader reform efforts can strengthen the perceived and actual connections between teacher quality, classroom instruction, and student learning.

In addressing teacher evaluation, state policymakers should consider taking action to:

- define teaching quality;
- focus evaluation policy on improving teaching practices;
- incorporate student learning into evaluation;
- create professional accountability
- train evaluators; and
- broaden participation in evaluation design.

Define teaching quality

States have defined academic standards for what every child needs to know. They also must clearly define what a highly qualified teacher needs to know and be able to do before they can purposefully construct a teacher evaluation policy.

To receive a teaching license, a candidate is required to hold a bachelor’s degree, maintain a minimum grade point average, and pass a criminal background check. In recent years, many states have established tougher licensure requirements. Most have set minimum scores for licensing exams, and many now require teacher-college graduates to have majored in an academic subject. Currently,

43 states require teaching candidates to pass at least a basic-skills test in order to receive a license⁵ or gain admission to a teacher preparation program.⁶

Traditional licensure does not guarantee teacher quality. For this reason, states are moving away from the notion of licensure as a way to ensure basic competence. They are designing performance-based licenses that require demonstration of subject knowledge and teaching skill, rather than basing licenses on course credits and hours of professional development. More than 30 states now require a candidate to pass Praxis II written tests of subject knowledge and pedagogy to receive a provisional teaching license. States are also beginning to use performance-based systems to create tiered professional designations—initial, provisional, professional, and master teacher licenses, for example.⁷

Performance-Based Standards for Colorado Teachers

Standard—*The teacher shall be knowledgeable about strategies, planning processes, assessment techniques, and appropriate accommodations to ensure student learning in a standards-based curriculum.*

Performance Criteria—*The teacher has demonstrated the ability to design short- and long-range standards-based instructional plans; develop valid and reliable assessment tools for the classroom; develop and utilize a variety of informal and formal assessments; assess, compare, and contrast the effects of various teaching strategies on individual student performance relative to content standards; use assessment data as a basis for standards-based instruction; provide effective verbal and written feedback that shape improvement in student performance on content standards; prepare students for [state assessments]; ensure that instruction is consistent with school district priorities and goals and [state content standards].*

A number of national efforts are underway to define what constitutes good teaching and to construct ways to measure it. More than 20 states have adopted standards, developed by the Interstate New Teacher Assessment and Support Consortium (INTASC), for beginning teachers. INTASC is currently translating those universal standards into discipline-specific standards. The National Board for Professional Teaching Standards has created standards and research-based assessments to identify accomplished teachers. Currently, 47 states and nearly 400 school districts provide fee reimbursements, salary supplements, or other financial incentives to encourage teachers to seek this distinction.⁸ These models for beginning and veteran teachers are good starting points for states that wish to create performance-based teaching standards. However, many states have already developed standards of their own (see Colorado and Kentucky examples).

Kentucky Performance Standards—Experienced Teacher Standards

Standard—*The teacher designs and plans instruction and learning climates that develop student abilities to use communication skills, apply core concepts, become self-sufficient individuals, become responsible team members, think and solve problems, and integrate knowledge.*

Performance Criteria—*The extent to which the teacher's plan focuses instruction on one or more of Kentucky's student academic expectations; develops the student's ability to apply knowledge, skills, and thinking processes; integrates skills, thinking processes, and content across disciplines; proposes learning experiences that challenge, motivate, and actively involve the learner; proposes learning experiences that are developmentally appropriate for learners and describes experiences for multiple levels of complexity to accommodate students at different levels of performance; incorporates strategies that address physical, social, and cultural diversity and show sensitivity to differences; establishes physical classroom environments to support the type of teaching and learning that is to occur; includes creative and appropriate uses of technology as a tool to enhance student learning; includes appropriate assessment strategies and processes; includes comprehensive and appropriate school and community resources that support learning; includes learning experiences that encourage students to be adaptable, flexible, resourceful, and creative.*

As states develop performance-based teaching standards and strengthen licensure requirements, they begin to specify the requisite knowledge and skills of a highly qualified teacher. States should then move forward and build evaluation policy around this definition. Without a clear framework about what teachers should know and be able to do at various stages in their careers, evaluation lacks clear purpose and fails to provide policymakers and school leaders with consistently reliable and useful information about what is happening in the classroom.

Focus evaluation policy on improving teaching practice

States should view evaluation as an informational tool to help administrators identify teachers who need additional or specialized assistance and to help individual teachers improve their instructional practices.

Evaluation typically serves as a process to measure job performance. However, it should also aim to improve teaching practice by providing constructive feedback to teachers and targeting professional development that is aligned with performance-based teaching standards. But few states have developed or made use of instruments—such as meaningful teacher assessments—to transform the traditional evaluation and licensure processes. Connecticut, North Carolina, and Ohio require teachers to demonstrate competent instructional practices in order to move beyond the initial license. In 2003, Ohio is slated to become the first state to fully implement Praxis III—a performance-based teacher assessment tool—to determine whether beginning educators can move from a provisional to a professional teaching license. States that develop or use such devices as part of the licensing process should be able to integrate them into evaluation as well.

Two teacher evaluation strategies hold promise in informing professional improvement: *peer review* and *portfolio assessment*. States can incorporate either into their overall teacher evaluation policy. Having grown out of peer assistance and mentoring, peer review permits consulting teachers to conduct formal evaluations of colleagues' performance and recommend professional assistance or even dismissal. To date, few school districts have implemented peer review programs. Chicago, Illinois, as well as Columbus and Toledo, Ohio, have among the best-known peer review programs. Chicago implemented its program as a way to improve the quality of teaching in its most troubled high schools. Columbus and Toledo target their peer review efforts at beginning teachers and those new to the district. In 1999, California enacted the first state peer review law.⁹

Portfolios are collections of a teacher's work, assembled to demonstrate command of subject content and instructional skill. They can be used for evaluative purposes but also to help teachers improve their instructional practices and increase their capacity for self-evaluation. Connecticut and Vermont are two states that use portfolio assessment within the teacher licensing process. In Connecticut, teachers must complete a highly structured, subject-specific portfolio assessment during their second year of teaching as a step in initial licensure.¹⁰ Vermont requires teaching candidates to assemble a portfolio as a centerpiece of their licensure applications and asks veteran teachers to construct portfolios again as part of the re-licensure process.¹¹

Incorporate student learning into evaluation

States should transform evaluation from a traditionally input-based process into an outcome-driven one. They should consider measurable student achievement as a principal outcome on which teachers are evaluated.

Increasingly, states are looking for ways to measure the effectiveness of teachers based, in part, on student learning. Twenty-four states are considering, developing, or implementing performance assessments of teachers—requiring demonstration of subject knowledge and pedagogy.¹² But only a handful consider measurable student achievement in evaluating (or re-licensing) veteran teachers—Delaware, Florida, Georgia, North Carolina, Tennessee, Texas, and Virginia.¹³ Tennessee uses assessment results to recommend professional development for teachers and allows, but does not require, value-added teacher-effect data to be used in evaluations; however, by law, those data are not publicly disclosed. Texas comes closer to making the connection between teacher quality and student achievement, basing one eighth of every teacher's evaluation on school-wide student performance.

One roadblock to evaluating teachers in this way is the lack of assessments, data systems, and evaluation processes capable of capturing the complexity of teaching skills—knowledge, pedagogy, classroom management—and their impact on student learning. As a result, most states and districts measure teacher performance based on input measures, such as professional development hours or completion of an induction program, rather than on student outcomes.¹⁴

Value-added assessment tracks the annual learning gains of students. It can gauge the impact an individual school or teacher has had on student achievement more accurately and fairly than traditional assessments, which measure cumulative achievement over time. Unlike percentiles, which are used to rank a student against his or her peers, a value-added approach determines a student's current level of attainment in a particular subject area. By aggregating scores by teacher, for instance, this type of assessment can be used to identify how much each teacher is influencing student learning. Only 16 states—notably Connecticut, Florida, North Carolina, Tennessee, and Texas—have developed assessments and data systems capable of enabling a value-added approach to student learning.¹⁵ Few are using that ability to fundamentally alter teacher evaluation, however.

Concerns about the reliability of test scores have influenced policymakers against holding teachers accountable for student learning. Student achievement reflects numerous influences unrelated to an individual teacher—past schooling, family background, behavioral and health problems, and other external factors. A single test may or may not properly reflect a student's true knowledge and skill. For these reasons, and because of the inherent volatility of small data samples, experts generally agree that teacher performance should never be based on the achievement of a single cadre of students; neither should it be based on student achievement alone.¹⁶

States therefore should broaden the variety of instruments and indicators used to evaluate teacher performance. Student tests—administered within a value-added assessment system—that are closely aligned with a state's academic standards should be one factor in evaluating teachers. Classroom observation, student work, teacher portfolios, self-evaluation, peer review, and verification of appropriate credentials potentially have roles to play as well.

Create professional accountability

Career ladders can provide states an opportunity to strengthen teacher evaluation policy and align it with performance-based teaching standards. Professional classifications (such as “beginning,” “mentor,” and “master” teacher) can also provide a framework through which to implement performance-based compensation.

Evaluations that determine teacher salary or influence a teacher's professional designation are more likely to affect instructional practices and teaching outcomes than ones that do not. Knowledge- and skill-based classification and pay-for-performance systems are revolutionary changes in the teaching profession, and they have yet to happen in most jurisdictions. Nonetheless, they can invest the evaluation process with greater purpose and allow teaching to have a more robust impact on student learning.

In 2001, Iowa piloted team-based, variable teacher pay in 10 school districts and 18 schools. The state is developing a new teacher evaluation system that would measure teaching performance.¹⁷ The new system would place teachers in one of four career classifications—*Beginning*, *Career*, *Career II*, or *Advanced*—providing for a comprehensive evaluation of beginning teachers at the end of their second year, a performance review of veteran teachers every three years, and an annual review of each teacher's individual development plan to help target professional development.¹⁸ Teachers would be assessed on professional standards that include content knowledge, classroom management, and an ability to enhance academic performance.¹⁹ The Beginning and Career I levels are in place, and the state is about to approve a pilot program for the Career II and Advanced classifications.²⁰ These developments could eventually result in the first statewide teacher pay-for-performance system.

The Milken Family Foundation has piloted a teacher performance-based accountability model in five Arizona school districts for the past two years (and more recently in Arkansas, Colorado, Florida, and South Carolina). Milken's Teacher Advancement Program is differentiated into three career classifications based on an individual's qualifications and responsibilities—*career*, *mentor*, and *master* teacher. Advancement is determined by three equal components: teaching processes (observable and documented instructional skills, teacher responsibilities, and knowledge); teaching

product (student achievement gains attributed to the teacher); and school product (student achievement gains attributed to the school).²¹

Train evaluators

Evaluators need preservice training opportunities to conduct more accurate and effective teacher assessments. Training might focus on skills such as analyzing effective teaching practice, determining a teacher's impact on student learning, and providing leadership for professional development and remedial assistance.

Cultivating and preparing school leaders for tasks such as teacher evaluation can enhance the quality of instruction, promote student achievement, and make evaluations more accurate and more consistent. MassPartners for Public Schools, a consortium of Massachusetts education organizations, argues that the main responsibility for teacher evaluation should not rest on the shoulders of individual principals. Instead, principals can more effectively influence evaluation by serving as instructional leaders and creating a professional community at their schools.²² Instructional leadership should be shared with teachers to promote what Harvard University's Richard Elmore calls "distributed leadership."²³ This may require greater investment in school-based professional development for school leaders, advanced principal and teacher leadership networks, and revamped leadership training within principal- and teacher-preparation programs.²⁴

The Iowa Educator Approval Training Program is one example of a state initiative aimed at improving the capability of evaluators to assess the performance of classroom teachers. Eligible evaluators include current and retired administrators, Area Education Agency staff, and National Board-certified teachers. The stated purpose of the program is to improve school district evaluators' skill in making employment decisions, making recommendations for beginning-teacher licensure, and providing support for moving teachers through career paths. The evaluation process supports individual teachers' career development plans, based upon the Iowa Teaching Standards and other established criteria.²⁵

Broaden participation in evaluation design

Policymakers must reach out to all education stakeholders, including teachers and administrators, to design a teacher evaluation system. Educators and school officials must have confidence in and an understanding of evaluation—prior to and during implementation—to ensure its long-term sustainability.

A viable evaluation system must first gain the acceptance of classroom teachers. Many states and localities with performance-based evaluation systems involved teachers in their design. One of the guiding principles outlined in the Missouri Department of Elementary and Secondary Education teacher evaluation guidelines calls on teachers and administrators to jointly design the system.²⁶ Teachers also must have the confidence to value the validity of assessment results and the knowledge about how they can use that information to improve their individual instructional practices.

It is often a much greater challenge to use an evaluation system as a gateway to additional reforms. The recent experience of the Cincinnati, Ohio, school district provides an important lesson. The district successfully designed a teacher evaluation system in cooperation with the Cincinnati Federation of Teachers. Annual evaluations measure 17 criteria, such as content knowledge and whether classroom tests are aligned with curriculum standards. They rate teachers at one of five performance levels: *apprentice*, *novice*, *career*, *advanced*, and *accomplished*.²⁷ However, in May 2002, Cincinnati teachers rejected an effort to tie compensation to teacher evaluations after the union urged its membership to vote against the proposal. That occurred despite a district study that found lower evaluation ratings for teachers translated into lower student performance and that showed lower-rated teachers were not negatively affected because they taught in the district's lowest-performing schools; indeed, they were distributed throughout the district.²⁸ The defeat of the Cincinnati pay-for-performance initiative underscores the importance of working with stakeholders and gaining buy-in for any proposed reform before attempted implementation.

Conclusion

Transforming a teacher evaluation system from one that assesses the job performance of teachers as a personnel action to one capable of improving teaching practice is a formidable challenge for state policymakers. It involves defining what teaching quality is, developing assessments to measure it, training effective evaluators, and capturing the impact of teaching on student learning. Fully realized, this approach also enables controversial reforms, such as pay-for-performance, that most jurisdictions have thus far resisted.

Strengthening teacher evaluation is a worthwhile challenge for Governors and state policymakers to undertake. It holds promise not only to professionalize teaching, but also to invest educators with greater information, confidence, and ability to improve their instructional practices and to help students achieve their fullest potential.

This *Issue Brief* was written by Liam Goldrick of the Education Policy Studies Division, National Governors Association Center for Best Practices, with funding from the Carnegie Corporation of New York.

¹ Linda Darling-Hammond and Deborah Loewenberg Ball, *Teaching for High Standards: What Policymakers Need to Know and Be Able to Do* (Philadelphia, Pa.: Consortium for Policy Research in Education, 1998); Ron Ferguson and Helen Ladd, "How and Why Money Matters: An Analysis of Alabama Schools," in *Holding Schools Accountable: Performance-Based Reform in Education* (Washington, D.C.: Brookings Institution, 1996), 265–298; William L. Sanders, S. Paul Wright, and Sandra Horn, "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation," in *Journal of Personnel Evaluation in Education* (Boston, Mass.: Kluwer Academic Press, 1997), 57–67.

² Charlotte Danielson, "New Trends in Teacher Evaluation," in *Educational Leadership* (Alexandria, Va.: Association for Supervision and Curriculum Development, February 2001), 14.

³ MassPartners for Public Schools, *Teaching Matters: Strengthening Teacher Evaluation in Massachusetts* (Monson, Mass.: MassPartners for Public Schools, February 2002), 4.

⁴ Allan Odden, *New Directions in Teacher Evaluation and Compensation* (PowerPoint presentation at the Second Annual Consortium for Policy Research in Education (CPRE) Annual Conference, Chicago, Ill., November 29–30, 2001), slide 17, <http://www.wcer.wisc.edu/cpre/conference/Nov01/allan.asp>.

⁵ *Quality Counts 2002* (Bethesda, Md.: Education Week, January 10, 2002), 78–79.

⁶ U.S. Department of Education, Office of Postsecondary Education, *Meeting the Highly Qualified Teachers Challenge: The Secretary's Annual Report on Teacher Quality* (Washington, D.C.: U.S. Department of Education, 2002), 54–55.

⁷ Eric Hirsch, Julia Koppich, and Michael Knapp, *Revisiting What State Are Doing to Improve the Quality of Teaching: An Update on Patterns and Trends* (Seattle, Wash.: Center for the Study of Teaching and Policy, University of Washington, February 2001), 16.

⁸ Bridget Curran, *States See Dramatic Increase in Number of National Board Certified Teachers* (Washington, D.C.: National Governors Association, December 6, 2001), www.nga.org/center/frontAndCenter/1,1188,C_FRONT_CENTER^D_2924,00.html.

⁹ National Governors Association (NGA), *Exploring Teacher Peer Review* (Washington, D.C.: NGA, January 12, 2000).

¹⁰ Suzanne M. Wilson, Linda Darling-Hammond, and Barnett Berry, *A Case of Successful Teaching Policy: Connecticut's Long-Term Efforts to Improve Teaching and Learning* (Seattle, Wash.: University of Washington, Center for the Study of Teaching and Policy, 2001), 15–17.

¹¹ National Association of State Boards of Education (NASBE), *State Improvement Initiatives: Vermont's Portfolio-Centered Teacher Licensure Process* (Alexandria, Va.: NASBE, May 1999).

¹² Eric Hirsch and Shelby Samuels, *Teaching in Vermont: An Inventory of Policies and Practices* (Denver, Colo.: National Conference of State Legislatures, 2001), 100.

¹³ Chester E. Finn, Jr., Marci Kanstoroom, and Michael J. Petrilli, *The Quest for Better Teachers: Grading the States* (Washington, D.C.: Thomas B. Fordham Foundation, 1999), 44; Lynn Cornett and Gale Gaines, *Performance Pay for Teachers: What Works and What Doesn't?* (Atlanta, Ga.: Southern Regional Education Board, August 2002); Dr. James H. Stronge, Heritage Professor of Educational Policy, College of William and Mary, e-mail to author, August 16, 2002; Education Commission of the States, *Recent State Legislation: Teacher Evaluation*, www.ecs.org.

¹⁴ Hirsch, Koppich, and Knapp, 15.

¹⁵ Education Commission of the States, *Teaching Quality>Certification/Licensure>What States Are Doing*, <http://www.ecs.org/html/IssueSection.asp?issueid=129&subissueid=79&s=What+States+Are+Doing>.

¹⁶ Thomas J. Kane, Douglas O. Staiger, and Jeffrey Geppert, "Randomly Accountable," in *Education Next* (Cambridge, Mass.: Hoover Institution, Spring 2002), 57–61.

¹⁷ Julie Blair, "Iowa Approves Performance Pay for Its Teachers," in *Education Week* (Bethesda, Md: Editorial Projects in Education, Inc., May 16, 2001).

¹⁸ Iowa State Education Association, *Evaluation Component: Teacher Quality/Compensation Legislation*, http://www.isea.org/whatsnew/teachercomp_items/eval_component.doc.

¹⁹ Sabrina W.M. Laine, *NCREL's Learning Point—Pay for Performance: New Thinking in Iowa* (Naperville, Ill.: North Central Regional Educational Laboratory, Spring 2002).

²⁰ Stephanie Pickens, Office of the Governor (Iowa), e-mail to author, November 7, 2002.

²¹ John Schacter, *Teacher Performance-Based Accountability: Why, What and How* (Santa Monica, Calif.: Milken Family Foundation, June 11, 2001), 9.

²² MassPartners for Public Schools, 11.

²³ Richard Elmore, *Building a New Structure for School Leadership* (New York, N.Y.: Albert Shanker Institute, 2000).

²⁴ National Staff Development Council, *Learning to Lead, Leading to Learn* (Oxford, Ohio: National Staff Development Council, December 2000).

²⁵ Jeff Berger, "Updated Information Regarding the Revised Teacher Evaluator Approval Training Program Development," (Des Moines, Iowa: Iowa Department of Education, March 5, 2002).

²⁶ Missouri Department of Elementary and Secondary Education, *Guidelines for Performance-Based Teacher Evaluation* (Jefferson City, MO: Missouri Department of Education, 1999).

²⁷ Cincinnati (Ohio) Public Schools, www.cpsboe.k12.oh.us/general/TchgProf/TES/TESTitle.html.

²⁸ Cincinnati (Ohio) Public Schools, "CPS Study Links Teacher Evaluation System to Student Achievement," <http://www.cpsboe.k12.oh.us/general/tchgprof/TES/TESstudy.html>.